Applied Probability Trust (October 11, 2000)

FOUR INTERESTING PROBLEMS CONCERNING MARKOVIAN SHAPE SEQUENCES

RICHARD COWAN,* *University of Sydney*

FRANCIS K.C. CHEN,** *University of Hong Kong*

Download of paper which appeared in *Adv. Appl. Prob.* **32**, 954-968 (1999).

Abstract

In earlier work, we investigated the dynamics of shape when rectangles are split into two. Further exploration, into the more general issues of Markovian sequences of rectangular shapes, has identified four particularly appealing problems. These problems, which lead to interesting invariant distributions on [0,1], have motivating links with the classical works of Blaschke, Crofton, D.G.Kendall, Rényi and Sulanke.

INVARIANT DISTRIBUTIONS; SHAPE; STOCHASTIC GEOMETRY; DISTRIBUTION THEORY; MARKOV PROCESSES; DYNAMICAL SYSTEMS; BROWNIAN MOTION.

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60D05
SECONDARY 60E05;60J05

0. Introduction

In an earlier paper, Chen and Cowan [2], we have explored Markovian sequences of rectangular shapes. The transition rules in that paper used the device of randomly splitting a given rectangle into two sub-rectangles and selecting one of these as the next in sequence. Interesting invariant distributions arose in that context. Some of those results were also presented in a higher-dimensional setting in Cowan [3].

Shape of a rectangle in $\mathbb{R}^2$ is defined for our purposes as the ratio of the shorter to longer side-lengths and so the distributions of shape have support $[0, 1]$. In Chen and Cowan, all invariant distributions had densities over this range.

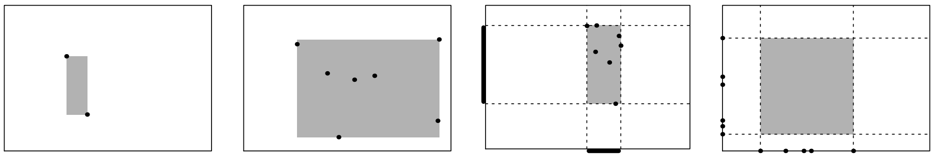
We have made an extensive study of invariant distributions for other transition rules which create Markovian shape sequences of rectangles. In the course of this study, four problems have appealed to us greatly. These problems, which we now describe, arise from natural geometric constructions and lead to interesting invariant distributions. Each problem has a motivating link with classical work, so we name the problems accordingly.

Our work also has a general affinity with the studies on Markovian shape sequences for triangles, as in Mannion [5, 6].

* Postal address: School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia.
** Postal address: Department of Statistics, University of Hong Kong, Pokfulam Road, Hong Kong.

1. Reminiscent of Rényi and Sulanke

Suppose at stage t, we have a rectangle R_t. If k independent and uniformly distributed points are placed in the interior of R_t, one might ask how these points could define a new rectangle, R_{t+1}. A natural method is one where R_{t+1} is the rectangle of smallest area having sides parallel to those of R_t and containing all k points [Figures 1(a), 1(b)]. This is a type of constrained *rectangular hull* of the k points, reminiscent of the *convex hull* construction in the famous Rényi-Sulanke [8] problem. The whole procedure is repeated for R_{t+1} and so on, thereby creating a random sequence of rectangles.



(a) (b) (c) (d)

Figure 1: The rectangular convex hull of k random points is shown as a shaded set in (a) with $k = 2$ or (b) with $k = 7$. In (c), $k = 7$ and the projected ranges of the k points onto the sides are shown as heavy lines. In (d), $n = 5$ random points on the long side and $m = 6$ random points on the short side generate the next rectangle as the Cartesian product of their ranges.

Preliminary calculations with the case $k = 2$, in Figure 1(a), suggested to us that the shapes of rectangles in this sequence degenerate. Progressively the rectangles become arbitrarily thin and their shapes converge in probability to zero. Suspecting that this conclusion may hold for all k, we imbedded the problem in a wider class where we expected some non-degeneracy would occur. Firstly we viewed the problem by projecting the k points onto both the short and long sides of R_t. These projected points form independent uniformly distributed points on each of these two sides. R_{t+1} is thus the Cartesian product of the *ranges* of the projected points on the long and short sides [Figure 1(c)].

Secondly, we generalised this range-generation method by placing n points uniformly distributed on the long side of R_t and m, distributed in a similar fashion, on the short side $(n, m \geq 2)$. R_{t+1} is defined as the Cartesian product of the ranges of these two, *possibly unequal*, samples of points on the two sides [Figure 1(d)]. The original idea of generating interior points is disregarded, unless $n = m = k$.

If Y is the shape of R_t and X that of R_{t+1}, one can find the transition density as

$$
\begin{aligned}
p(x|y) &= \frac{m(m-1)n(n-1)x^{n-2}}{(m+n-1)y^m}\left[\frac{x^{m-n}y + y^{m+n-1}}{m+n-2} - \frac{x^{m-n+1} + xy^{n+m}}{m+n}\right] && 0 < x < y, \\
&= \frac{m(m-1)n(n-1)y^{n-1}}{(m+n-1)x^{n+1}}\left[\frac{x(1+x^{2n-2})}{m+n-2} - \frac{y(1+x^{2n})}{m+n}\right] && y < x < 1.
\end{aligned}
$$

If f is an invariant probability density function of shape, then the standard equation,

$f(x)=\int_0^1 p(x|y)f(y)dy$, becomes

$$
\begin{aligned}
(1)\qquad f(x) =& \frac{m(m-1)n(n-1)}{m+n-1}\Bigg[\int_0^1\left(\frac{x^{n-2}y^{n-1}}{m+n-2}-\frac{x^{n-1}y^n}{m+n}\right)f(y)dy\\
&+\int_0^x\left(\frac{y^{n-1}}{x^n(m+n-2)}-\frac{y^n}{x^{n+1}(m+n)}\right)f(y)dy\\
&+\int_x^1\left(\frac{x^{m-2}}{y^{m-1}(m+n-2)}-\frac{x^{m-1}}{y^m(m+n)}\right)f(y)dy\Bigg].
\end{aligned}
$$

This integral equation can be converted to a differential equation which is of seventh order and of Euler type. Equations of Euler form have particular solutions of the form x^r where in our case r equals either $m-n-1, n-2, n-1, 0, -1$ or the roots of $r^2+r(n+2-m)-2n(m-1)=0$. Since $(n+2-m)^2+8n(m-1)=(m+n)^2+4(m-1)(n-1)>(m+n)^2$ under our condition that $n,m\geq 2$, the roots of this quadratic equation are both real. The larger, c say, is

$$
(2)\qquad c := \tfrac{1}{2}[m-n-2+\sqrt{(m+n)^2+4(m-1)(n-1)}] > m-1.
$$

A similar argument shows that the smaller root, c' say, is less than $-(n+1)$. Note that $c\sim m$ as $m\to\infty$ for each fixed n.

Not all solutions of the differential equation need be solutions of (1); in particular, since f must integrate to 1 over the range $[0,1]$, a function containing a term of the form $x^r, r\leq -1$, cannot be the invariant solution we seek. Thus -1 and c' are eliminated, as is $m-n-1$ if $m\leq n$. So, provided $m>n$ and the (m,n)-choice does not lead to multiplicity of those r values which are greater than -1,

$$
(3)\qquad f(x) = Ax^0+Bx^{n-1}+Cx^{n-2}+Dx^{m-n-1}+Ex^c
$$

is a candidate solution of (1), with no other form being possible. Subtract $f(x)$ from both sides of (1) and substitute (3). This gives a lengthy expression involving terms in $x^{n-1}, x^{n-2}, x^{m-1}$ and x^{m-2} together with a constant which can only equal zero if $A=0$. Terms in x^{m-n-1} and x^c appear briefly in the calculation too, but cancel out. After A is set to zero, we equate the four remaining coefficients to zero, yielding four homogeneous linear equations for B, C, D and E. The 4 equations have rank of 3 only, leaving open the possibility of a non-zero solution. Because of the added condition that f integrate to 1, we obtain a fifth linear equation, linearly independent of the others, and hence a non-homogeneous system of full rank. Thus we establish, in this uncomplicated (m,n)-context, that, for the f in (3) to both satisfy (1) and integrate to 1, it is necessary and sufficient that A=0 and that

$$
\begin{aligned}
B &= -\frac{(m-n)(m-n-1)[n(m-1)-c]K}{m(m-2n)(1+c-n)},\\
C &= \frac{n(m-n)(m-n+1)K}{(1+m-2n)(2+c-n)},\\
D &= \frac{n(n-1)[n(2n-1)-(m+1-2n)c]K}{m(1+m-2n)(1+c-m+n)(m-2n)},\\
E &= \frac{(m-n)(2+c-m)(1+c-m)(c+n)K}{m(1+c-n)(2+c-n)(1+c-m+n)}, \quad \text{where}\\
K &:= \frac{n(n-1)(1+c)(m-n)}{(m-1)[n(m+n-1)+(n-1)c]}
\end{aligned}
\tag{4}
$$

and where c is given by (2). (Alternative arrangements for the expressions above, for example, eliminating the radical contained within c from the denominator, are less compact.) As an example of the general result, consider $(m, n) = (7, 3)$. Then $c = 1 + \sqrt{37}$ and our invariant density, illustrated in Figure 2 along with other cases, is

$$f(x) = x\,\left(\tfrac{40\,(925-16\sqrt{37})}{8547} - \tfrac{16\,(20-\sqrt{37})}{77}\,x + \tfrac{4\,(433-64\sqrt{37})}{2541}\,x^2 + \tfrac{16\,(218\sqrt{37}-851)}{94017}\,x^{\sqrt{37}}\right).$$

Note that, as $m \to \infty$,

$$f(x) \to \frac{n(n-1)x^{n-2}[n-(n-1)x]}{(2n-1)}, \tag{5}$$

such forms being depicted as dashed curves in Figure 2 and as a family in Figure 3(a).

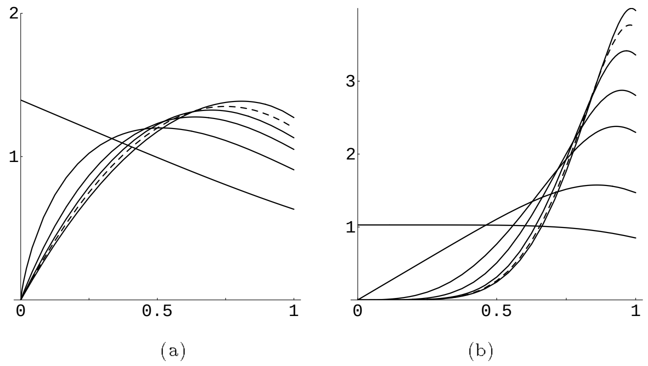


Figure 2: (a) The invariant densities for $n = 3$ and $m \in \{4, 5, 6, 7, 18\}$ are shown as solid curves, with $f(1)$ increasing as m increases within this range. For $m > 18$, $f(1)$ decreases with m, finally yielding the asymptotic case shown as a dashed curve. (b) The case $n = 8$ with $m \in \{9, 10, 12, 14, 18, 53\}$, with similar increasing behaviour for $f(1)$. For $m > 53$, $f(1)$ decreases. The dashed curve is the limiting case.

We now clarify the situation when $m \leq n$, concluding as anticipated initially that there is degeneracy of shapes in these cases.

Theorem 1. *When $m \leq n$, the only solution of the integral equation (1) is $f(x) = 0$, and so, there does not exist an invariant probability density function.*

Proof. Let $m \leq n$. The function f in (3) is still a contender for the solution of (1) provided D is set equal to zero (*except* when $m = n = 2$ because a multiplicity occurs in the four remaining exponents only in that exceptional case). If $n > 2$, it is necessary to consider many separate cases. $D = 0$ throughout.

(i) $2 < m < n-1$. As before we find that $A = 0$ and obtain four homogeneous linear equations of rank 3, but now in only 3 variables, B, C and E. So we conclude that $B = C = E = 0$ and thus there is no solution to (1) other than the zero function.
(ii) $2 < m = n-1$. Again $A = 0$ but note that the third integral in (1) yields terms in $x^{n-2}\log(x)$; these equal zero iff $C = 0$. $B = E = 0$ follows from equating the coefficients of $x^{n-1}, x^{n-2}(\equiv x^{m-1})$ and $x^{n-3}(\equiv x^{m-2})$ to zero, the equations being of rank 2.
(iii) $2 = m < n-1$. Now the third integral yields $\log(x)$ terms which are zero iff $A = 0$, as well as terms in $x^0(\equiv x^{m-2}), x^1(\equiv x^{m-1}), x^{n-1}$ and x^{n-2}, the coefficients of which form 4 equations of rank three in B, C, E.
(iv) $2 = m = n-1$. Now $n = 3$. Terms arise in $\log(x)$ these being zero iff $A = 0$. Others in $x\log(x)$ are zero iff $C = 0$. Equations of rank 2 from the coefficients of x-powers $0(\equiv m-2), 1(\equiv m-1 \equiv n-2)$ and $2(\equiv n-1)$ prove that $B = E = 0$.
(v) $2 < m = n$. Here $A = 0$ by consideration of the constant, but now the third integral produces terms in $x^{n-2}\log(x)$ which equal zero iff $C = 0$ and some in $x^{n-1}\log(x)$, zero iff $B = 0$. $E = 0$ follows the coefficients of $x^{n-1}(\equiv x^{m-1})$ and $x^{n-2}(\equiv x^{m-2})$ terms (equations being rank 1).

We now deal with the multiplicity that occurs when $m = n = 2$. Standard theory says that $f(x)$ must be of the form $A + Bx + C\log(x) + Ex^c$. Substitution in (1) yields a $[\log(x)]^2$ term which is zero iff $C = 0$, an $x\log(x)$ term which is zero iff $B = 0$ and $\log(x)$ terms which imply $A = 0$. $E = 0$ follows readily. $\square$

Theorem 1 seems very reasonable. Intuitively, in order to avoid a trend toward thinner rectangles, one must take a larger proportion of the short side, than of the long side, when constructing the next rectangle in the sequence. This is achieved by having $m > n$.

For these cases where $m > n$, we find that an analysis of multiplicities is needed. Only when $n > 2, m \neq n+1, m \neq 2n$ and $m \neq 2n-1$ are there no multiplicities. When these conditions are violated, the analysis leading to the solution given by (3) and (4) is invalid. Let us deal firstly with the exceptional cases when $n > 2$.

When $n > 2, m = 2n$, there is a multiplicity of 2 at $r = n-1$. Standard theory suggests that the solution should be $f(x) = A + Bx^{n-1} + Cx^{n-2} + Dx^{n-1}\log(x) + Ex^c$.

Detailed analysis following our now established methods shows that $A = 0$ and

$$
\begin{aligned}
B &= \frac{(1+c)(c-n)(n-1)nH}{4(2n-1)(1+c-n)^3(c+n)} \\
C &= \frac{(1+c)(n-1)n^4(n+1)}{(2n-1)(2+c-n)(1+c-n)(c+n)} \\
D &= \frac{(1+c)(n-1)^2n^3(c-n)}{2(2+c-n)(1+c-n)^2(c+n)} \\
E &= \frac{(1+c)(1+c-2n)(n-1)n^4}{(2n-1)(2+c-n)(1+c-n)^3(c+n)},
\end{aligned}
\tag{6}
$$

where $H := c(n-1) - 2n[1 - 3n + 2(n+1)n^2]$.

When $n > 2, m = 2n - 1$, the multiplicity is now at $r = n - 2$ so $f(x) = A + Bx^{n-1} + Cx^{n-2} + Dx^{n-2}\log(x) + Ex^c$ is the candidate. We find that $A = 0$ and

$$
\begin{aligned}
B &= \frac{(1+c)(n-2)(n-1)^3n}{(2n-1)(2+c-n)(c-1+3n)} \\
C &= \frac{(1+c)(3+c-n)(1+c-n)nH}{4(2n-1)(2+c-n)^3(c-1+3n)} \\
D &= -\frac{(1+c)(n-1)^2n^2(3+c-n)}{2(2+c-n)^2(c-1+3n)} \\
E &= \frac{(1+c)(3+c-2n)(n-1)^3n}{(2n-1)(2+c-n)^3(c-1+3n)},
\end{aligned}
\tag{7}
$$

where now $H := nc - 2(2 - 10n + 15n^2 - 10n^3 + 2n^4)$.

When $n > 2, m = n + 1$, x^{m-n-1} collapses to x^0, so the only candidate solution is of the form $f(x) = D + Bx^{n-1} + Cx^{n-2} + A\log(x) + Ex^c$. The analysis shows that $A = 0$, and so the $\log(x)$ term does not remain in the solution. We also find the other four coefficients and note that they agree with (4), under the substitution $m = n+1$ (hence explaining our rather odd juxtaposition of A and D above).

We now deal with the cases when $n = 2$. When $n = 2, m \geq 5$, there is a multiplicity of 2 at $r = 0$. Standard theory suggests that the solution be $f(x) = C + Bx + A\log(x) + Dx^{m-3} + Ex^c$, but it turns out that $A = 0$ and an $(n = 2)$-version of (4) gives the other coefficients. Note that, as $m \to \infty$, $f(x) \to 2(2-x)/3$, agreeing with (5).

When $n = 2, m = 4$, the only candidate solution is $f(x) = C + Bx + A\log(x) + Dx\log(x) + Ex^c$ where $c = 2\sqrt{3}$. We find that $A = 0$ and, for $x \in (0, 1]$,

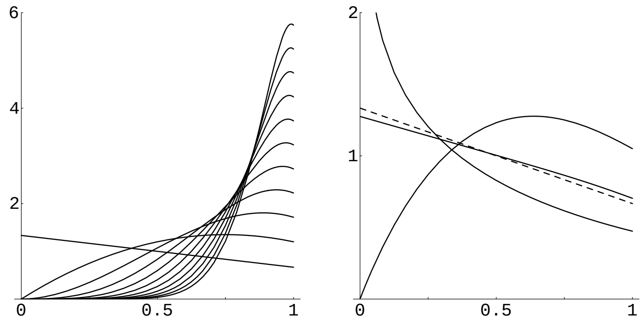
$$
f(x) = \tfrac{2\left(27-\sqrt{3}\right)}{33} - \tfrac{\left(4663-464\sqrt{3}\right)}{3993}\,x + \tfrac{2\left(69-16\sqrt{3}\right)}{363}\,x\,\log(x) + \tfrac{2\left(131\sqrt{3}-149\right)}{3993}\,x^{2\sqrt{3}},
$$

a result which is consistent with an $(n = 2)$-version of (6).

Finally, when $n = 2, m = 3$, there is a multiplicity of order 3 at $r = 0$. Thus we need to try $f(x) = D + Bx + C\log(x) + A[\log(x)]^2 + Ex^c$. The result, which is consistent with (7) when $n = 2$, is $A = B = 0$ and, for $x \in (0, 1]$,

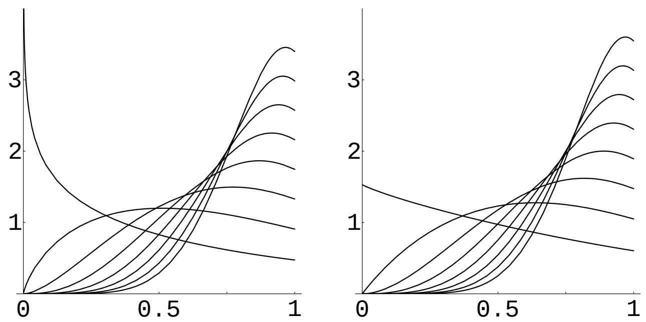
$$
f(x) = \tfrac{605+13\sqrt{33}}{1536} - \tfrac{111-\sqrt{33}}{192}\,\log(x) + \tfrac{19\sqrt{33}-61}{1536}\,x^{(\sqrt{33}-1)/2}.
$$

Figures 3(b) shows the $(n = 2)$-cases for a range of m whilst Figure 4 illustrates the invariant densities when $m = 2n$ or $2n - 1$ (as does Figure 2(a) in the $(n = 3)$-case).



(a) (b)

Figure 3: (a) Here the limiting form of the invariant density as $m \to \infty$ is shown for $n = 2, 3, \ldots, 12$; $f(1)$ increases monotonely with n. (b) Some invariant densities for $n = 2$. The solid curves show $m \in \{3, 4, 11\}$ whilst the dashed curve is the limiting case for large m. $f(0)$ is infinite for $m = 3$ and equals zero for $m = 4$. Had other cases been shown, the $(m = 5)$-curve would be imperceptibly different from the dashed curve, whilst as m increases, the curves would gradually shift to the position shown by $m = 11$ before gradually shifting back to the position of the dashed curve.



(a) (b)

Figure 4: The invariant densities for: (a) $m = 2n - 1$; (b) $m = 2n$. In both graphs $n \in \{2, 3, 4, 5, 6\}$, with $f(1)$ increasing as n increases.

2. In the spirit of Crofton

Given a rectangle R_t, draw an isotropic uniformly distributed (IUR) transect $\mathcal{T}$. An IUR transect is a line uniformly sampled from the space of all lines which intersect R_t. It was Crofton [4] who, last century, first introduced this concept. In our application, this transect will cut R_t into two polygons but these will not of course be rectangles. We create however a new rectangle R_{t+1}, the next in sequence, in a natural way. It is defined as the unique rectangular subset of R_t having (a) sides parallel to those of R_t and (b) $\mathcal{T} \cap R_t$ as a diagonal. We shall find that there is a very surprising fact about the shape of R_{t+1} and, given that Crofton's classical results contained an element of surprise for the mathematical world of his day, our result is truly in the spirit of Crofton.

Suppose that the shape of R_t is y. Let $\mathcal{A}$ be the event that $\mathcal{T}$ cuts *adjacent* sides of R_t, for example, as in Figure 5(b). Let $\mathcal{L}$ be the event that opposite *long* sides are cut [Figure 5(c)] and $\mathcal{S}$ the event that opposite *short* sides are cut [Figure 5(d)].

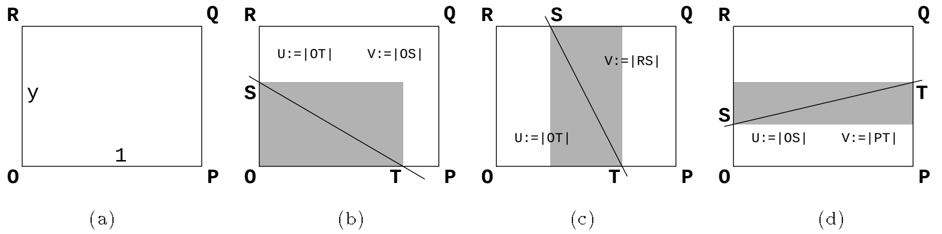


Figure 5: (a) A rectangle of shape y. (b-d) The three ways that an IUR transect can cut the rectangle and the resulting new rectangle shown shaded.

The event that $\mathcal{T}$ passes through a corner has probability zero, so $\mathbb{P}\{\mathcal{A}\}+\mathbb{P}\{\mathcal{L}\}+\mathbb{P}\{\mathcal{S}\}=1$. (Here, for brevity, we do not emphasise notationally the conditioning on y, though we shall see that y appears in most of our expressions below.) For some set D, let $\{\mathcal{T}\uparrow D\}:=\{\mathcal{T}\cap D\neq\emptyset\}$; we say that $\mathcal{T}$ "hits" D. We also define $\{\mathcal{T}\uparrow(D_1\&D_2)\}$ by $\{\mathcal{T}\uparrow D_1\}\cap\{\mathcal{T}\uparrow D_2\}$.

The line-segment RP [Figure 5(a)] is a convex subset of the rectangle $OPQR$ and so $\mathbb{P}\{\mathcal{T}\uparrow RP\}$ equals the perimeter of RP, namely $2\sqrt{1+y^2}$, divided by the perimeter of $OPQR$, namely $2(1+y)$. Since $\{\mathcal{T}\uparrow RP\}=\mathcal{L}\cup\mathcal{S}\cup\{\mathcal{T}\uparrow(OR\&RQ)\}\cup\{\mathcal{T}\uparrow(OP\&PQ)\}$, we have $\frac{1}{2}\mathbb{P}\{\mathcal{A}\}+\mathbb{P}\{\mathcal{L}\}+\mathbb{P}\{\mathcal{S}\}=\sqrt{1+y^2}/(1+y)$. Similarly, using OP instead of RP, one sees that $\frac{1}{2}\mathbb{P}\{\mathcal{A}\}+\mathbb{P}\{\mathcal{L}\}=1/(1+y)$. So, from these three equations, we have

$$(8)\quad \mathbb{P}\{\mathcal{A}\}=\frac{2(1+y-\sqrt{1+y^2})}{1+y},\quad \mathbb{P}\{\mathcal{L}\}=\frac{\sqrt{1+y^2}-y}{1+y},\quad \mathbb{P}\{\mathcal{S}\}=\frac{\sqrt{1+y^2}-1}{1+y}.$$

Define the variables U and V as in Figures 5(b-d), with V being measured on the rectangle's *short* side in Figure 5(b). For $u\in[0,1], v\in[0,y]$, using Figure 6(a) and (8),

$$\begin{aligned}
F_{\mathcal{A}}(u,v) :=&\mathbb{P}\{U\leq u, V\leq v,\mathcal{A}\}\\
=&4\,\mathbb{P}\{\mathcal{T}\not\uparrow ABRP,\mathcal{T}\uparrow(OR\&OP)\}\\
=&4(\mathbb{P}\{\mathcal{T}\not\uparrow ABRP\}-\mathbb{P}\{\mathcal{T}\not\uparrow ABRP,\{\mathcal{T}\uparrow(OR\&OP)\}'\})\\
=&4(1-\frac{perim(ABRP)}{perim(OPQR)}-\frac{\mathbb{P}\{\mathcal{A}\}}{4})\\
=&4(1-\frac{y-v+1-u+\sqrt{u^2+v^2}+\sqrt{1+y^2}}{2(1+y)}-\frac{1+y-\sqrt{1+y^2}}{2(1+y)})\\
=&\frac{2(u+v-\sqrt{u^2+v^2})}{1+y}.
\end{aligned}$$

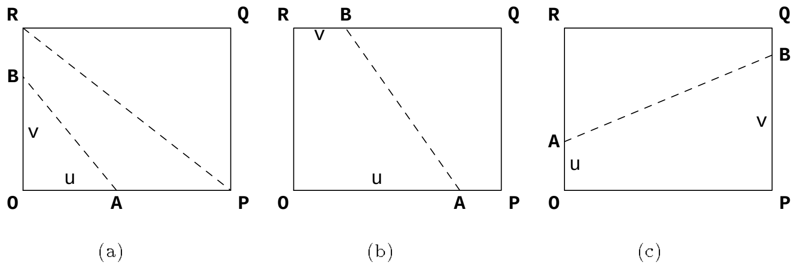


Figure 6: These three drawings illustrate an argument in the text.

For $u, v \in [0,1]$ and using Figure 6(b),

$$\begin{aligned}
F_{\mathcal{L}}(u,v) :=& \mathbb{P}\{U \leq u, V \leq v, \mathcal{L}\} \\
=& \mathbb{P}\{\mathcal{T} \not\uparrow BQPA, \mathcal{L}\} \\
=& \mathbb{P}\{\mathcal{T} \not\uparrow BQPA\} - \mathbb{P}\{\mathcal{T} \not\uparrow BQPA, \mathcal{L}'\} \\
=& 1 - \frac{perim(BQPA)}{perim(OPQR)} - \mathbb{P}\{\mathcal{T} \uparrow (OR\&RB)\} - \mathbb{P}\{\mathcal{T} \uparrow (OR\&OA)\} \\
=& 1 - \frac{y - v + 2 - u + \sqrt{y^2 + (u-v)^2}}{2(1+y)} - \frac{F_{\mathcal{A}}(u,y) + F_{\mathcal{A}}(v,y)}{4} \\
=& \frac{\sqrt{u^2+y^2} + \sqrt{v^2+y^2} - y - \sqrt{y^2+(u-v)^2}}{2(1+y)}.
\end{aligned}$$

Similarly, for $u, v \in [0, y]$, using Figure 6(c),

$$\begin{aligned}
F_{\mathcal{S}}(u,v) :=& \mathbb{P}\{U \leq u, V \leq v, \mathcal{S}\} \\
=& 1 - \frac{2y - u - v + 1 + \sqrt{1 + (u-v)^2}}{2(1+y)} - \frac{F_{\mathcal{A}}(1,u) + F_{\mathcal{A}}(1,v)}{4} \\
=& \frac{\sqrt{1+u^2} + \sqrt{1+v^2} - 1 - \sqrt{1+(u-v)^2}}{2(1+y)}.
\end{aligned}$$

If the shape of R_{t+1} is X, then $\mathbb{P}\{X > x\} = \mathbb{P}\{X > x, \mathcal{A}\} + \mathbb{P}\{X > x, \mathcal{L}\} + \mathbb{P}\{X > x, \mathcal{S}\}$. We now analyse these three terms. For $x \in [0,1]$,

$$\begin{aligned}
\mathbb{P}\{X > x, \mathcal{A}\} =& \mathbb{P}\{\min\left(\frac{V}{U}, \frac{U}{V}\right) > x, \mathcal{A}\} \\
=& \mathbb{P}\{xU < V < U/x, \mathcal{A}\} \\
=& \int_0^{xy} \int_{xu}^{u/x} f_{\mathcal{A}}(u,v)\,dv\,du + \int_{xy}^{1} \int_{xu}^{y} f_{\mathcal{A}}(u,v)\,dv\,du \qquad 0 \leq x < y \\
=& \int_0^{xy} \int_{xu}^{u/x} f_{\mathcal{A}}(u,v)\,dv\,du + \int_{xy}^{y/x} \int_{xu}^{y} f_{\mathcal{A}}(u,v)\,dv\,du \quad y \leq x \leq 1,
\end{aligned} \tag{9}$$

where we have introduced one of the *improper densities*, listed below.

$$\begin{aligned} f_{\mathcal{A}}(u,v) &:= \frac{\partial^2 F_{\mathcal{A}}(u,v)}{\partial u \partial v} = \frac{2uv}{(u^2+v^2)^{3/2}(1+y)} \\ f_{\mathcal{L}}(u,v) &:= \frac{\partial^2 F_{\mathcal{L}}(u,v)}{\partial u \partial v} = \frac{y^2}{2(y^2+(u-v)^2)^{3/2}(1+y)} \\ f_{\mathcal{S}}(u,v) &:= \frac{\partial^2 F_{\mathcal{S}}(u,v)}{\partial u \partial v} = \frac{1}{2(1+(u-v)^2)^{3/2}(1+y)}. \end{aligned}$$

Evaluation of (9) yields

$$\begin{aligned} \mathbb{P}\{X > x, \mathcal{A}\} &= \frac{2(1+y-\sqrt{1+x^2}\sqrt{1+y^2})}{(1+y)\sqrt{1+x^2}} && x < y, \\ &= \frac{2y(1-x)}{(1+y)\sqrt{1+x^2}} && x > y. \end{aligned} \tag{10}$$

To evaluate $\mathbb{P}\{X > x, \mathcal{L}\}$, we define $W := |U - V|$ when $\mathcal{L}$ or $\mathcal{S}$ occurs and note that for $w \in [0,1]$,

$$\begin{aligned} \mathbb{P}\{W \le w, \mathcal{L}\} &= 2\left(\int_0^{1-w}\int_u^{u+w} f_{\mathcal{L}}(u,v)\,dv\,du + \int_{1-w}^1\int_u^1 f_{\mathcal{L}}(u,v)\,dv\,du\right) \\ &= \frac{w+y^2-y\sqrt{w^2+y^2}}{(1+y)\sqrt{w^2+x^2}}. \end{aligned}$$

Thus for $x \in [0,1]$,

$$\begin{aligned} \mathbb{P}\{X > x, \mathcal{L}\} &= \mathbb{P}\{\min\left(\frac{y}{W}, \frac{W}{y}\right) > x, \mathcal{L}\} \\ &= \mathbb{P}\{xy < W < y/x, \mathcal{L}\} \\ &= \frac{\sqrt{1+x^2}\sqrt{1+y^2}-x-y}{(1+y)\sqrt{1+x^2}} && x < y \\ &= \frac{(1-x)(1-y)}{(1+y)\sqrt{1+x^2}} && x > y. \end{aligned} \tag{11}$$

Similarly, for $w \in [0, y]$,

$$\begin{aligned} \mathbb{P}\{W \le w, \mathcal{S}\} &= \frac{1+wy-\sqrt{1+w^2}}{(1+y)\sqrt{1+w^2}}, && \text{so,} \\ \mathbb{P}\{X > x, \mathcal{S}\} &= \mathbb{P}\{W > x, \mathcal{S}\} \\ &= \frac{\sqrt{1+x^2}\sqrt{1+y^2}-1-xy}{(1+y)\sqrt{1+x^2}} && x < y \\ &= 0 && x > y. \end{aligned} \tag{12}$$

Combining (10), (11) and (12), one obtains $\mathbb{P}\{X > x\}$. Differentiating yields the transition density function $p(\cdot|y)$. Here the dependence on y is emphasised for the first time in our argument, *but* there is a surprise! For all $y \in [0, 1]$,

$$p(x|y) = \frac{1+x}{(1+x^2)^{3/2}}. \tag{13}$$

The shape distribution of our new rectangle is actually not dependent on y, the previous shape! This means that, regardless of the shape of R_0, R_1 and all subsequent rectangles have shape pdf given by the RHS of (13); furthermore all shapes in the sequence are independent. A plot of (13) is shown in Figure 7(a).

Thus the Markov process reaches stationarity after just one step, a phenomenon rarely seen, though a referee has pointed out that this occurs in the Hardy-Weinberg law of mathematical genetics.

3. The link with Brownian motion following Kendall

Consider a point P, initially at the planar origin O. Now let P move as a standardised planar Brownian motion, that is, its cordinates (X_t, Y_t) at time t undergo independent standard one-dimensional Brownian motions. At each time t, we form a rectangle with sides parallel to the axes and having OP as a diagonal. How does the shape of this evolving rectangle change? Does it have a limiting distribution?

Our interest in these questions arises from Kendall's [7] work concerning the dynamics of shape for $n \geq 3$ randomly-diffusing points in the plane. His notion of shape is inapplicable, however, for $n < 3$; two points alone do not have a shape and neither does our single point, P. *Our* shape arises not from P itself, but from a geometric object formed from a diffusing P and three other points.

The rectangle's shape S_t at time t is given by

$$S_t = \min\left(\frac{|X_t|}{|Y_t|}, \frac{|Y_t|}{|X_t|}\right).$$

Because the Brownian motions have no drift and equal unit variances, we readily see that, for $s, t > 0$,

$$\mathbb{P}\{S_t > s\} = \mathbb{P}\{s < \frac{|X_t|}{|Y_t|} < \frac{1}{s}\} \tag{14}$$

$$= \mathbb{P}\{s < \frac{|X_t|/\sqrt{t}}{|Y_t|/\sqrt{t}} < \frac{1}{s}\}$$

$$= \int_s^{1/s} \frac{2}{\pi(1+z^2)} dz, \tag{15}$$

the last expression arising due to the well-known fact that the ratio of two standard normal variates is Cauchy distributed. (The restriction to absolute values gives the half-Cauchy pdf which is our integrand.) Thus we have that, for *all* time $t > 0$, S_t has the pdf on [0,1]

$$f(s) = \frac{4}{\pi(1+s^2)}. \tag{16}$$

Naturally this form, which is plotted in Figure 7(a), is also the limiting form.

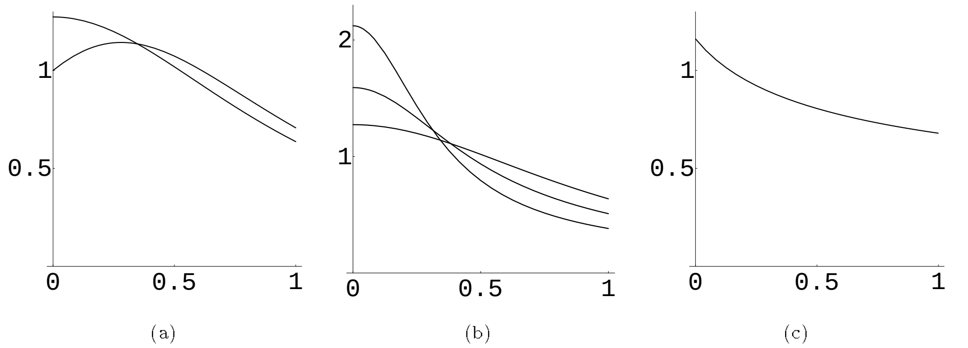


Figure 7: (a) The invariant densities for the "Crofton" diagonal-construction method and the standardised driftless Brownian-motion method, the former being higher at $x = 0$. (b) The driftless Brownian-motion case with unequal variances, the ratio b being 1, 2 or 3. (c) The Blaschke method, showing the density for the continuous component of the equilibrium distribution.

An interesting distribution arises if we relax the stipulation of equal unit variances in the x- and y-motions of P. Suppose that the ratio of one standard deviation to the other is $b > 0$. A minor variation in the foregoing argument (14, 15) leads to the distribution of S_t as

$$f(s) = \frac{2b(1+s^2)(1+b^2)}{\pi(s^2+b^2)(1+b^2s^2)}, \tag{17}$$

for all $t > 0$ [Figure 7(b)]. One can also show that starting P at a location away from the origin leads to the same limiting distributions (16) or (17).

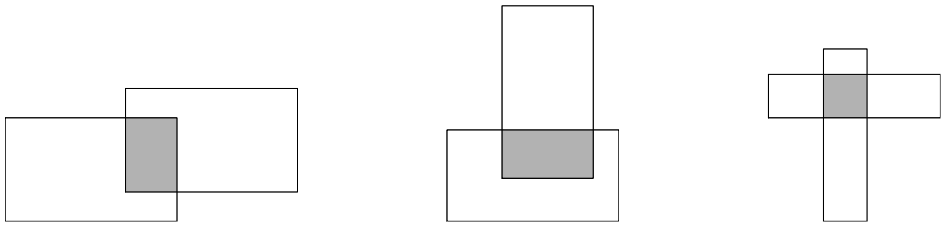
Relaxing the zero-mean assumption for the Brownian motion components markedly alters the situation, as one might expect. A detailed analysis shows that, if $X_t \sim N(x_0 + \mu t, \sigma^2 t), Y_t \sim N(y_0 + a\mu t, b^2\sigma^2 t)$, the shapes converge in distribution to a constant, namely the smaller of $|a|$ and $1/|a|$. In effect, P is drifting along the line $y = ax$ and in due course its distance from this line becomes arbitrarily small compared with the distance OP. For the case $a = 1$, the proof involves four double integrals to evaluate $\mathbb{P}\{S_t > s\}$ as given in (14), one for each quadrant in which P may lie at time t. Conversion to polar coordinates helps reduce these to single integrals. The integrands in each of these are bounded functions of t, for $t \geq 0$ and so interchange of limit and integration is permitted. One thus establishes, when $\mu \neq 0$, that $\mathbb{P}\{S_t > s\} \to 1$, for any $s \in [0, 1)$, and so, shapes converge in distribution to 1. When $a < 1$, we consider the problem of shape convergence when a point P' starts at $(x_0, y_0/a)$ and follows the path $(X_t, Y_t/a)$. In the early stages of this new process, the shape, S'_t say, may differ from S_t/a, but there is eventually a coupling between the two shape processes, that is, there is with probability one a time t_c such that, for all $t \geq t_c$, $S'_t = S_t/a$. Since $S'_t \to 1$, we see that $S_t \to a$. A similar argument applies for the case $a > 1$.

As a final remark, we note that if *two* opposite corners of a rectangle are allowed to move as Brownian motion, instead of just one in the problem above, one merely considers the relative motion of one corner to the other. Thus the same invariant distributions for rectangle shape, (16) and (17), arise.

4. A problem in the style of Blaschke

At time $t \in \{0, 1, 2, \ldots\}$, take a congruent copy of R_t and place it randomly in the plane in a manner such that the intersection of the copy with R_t itself is a rectangle; this becomes R_{t+1}. The randomness is uniform over all positions and orientations which achieve this rectangular intersection. There are, of course, only two orientations permitted for the copy, major axes either aligned with those of R_t or oriented orthogonally [Figure 8].

This notion of allowing a domain to undergo a uniformly distributed intersection with a congruent copy is highly reminiscent of the classical geometric work of Blaschke [1].



(a) (b) (c)

Figure 8: Blaschke-style hitting of the rectangle with its congruent copy.

Before rushing in to establish an invariant probability density of shape, we take note of Figure 8(c), which illustrates that the square shape, $X_{t+1} = 1$, has positive probability of arising from any non-square predecessor. Indeed, the smaller is the shape X_t, the higher is the chance that $X_{t+1} = 1$. Thus we can anticipate that any invariant distribution of shape will have a probability mass at 1. No density will exist and we need to work with distribution functions.

We find, after detailed calculation that the transition distribution function of shape X given its predecessor is y is

$$\begin{aligned} P(x|y) :=& \mathbb{P}\{X \leq x|y\} \\ =& \frac{2x(1+y)^2}{1+6y+y^2} && x < y \\ =& \frac{2y(2x^2 + x^2y + 2x - y)}{x(1+6y+y^2)} && x > y. \end{aligned}$$

Note that $P(1-|y) = 8y/(1+6y+y^2)$. So $\mathbb{P}\{X = 1|y\} = (1-y)^2/(1+6y+y^2)$. Note also that $P(x|1) = x$.

Assume that an invariant distribution exists and denote it by F. Introduce $a := 1 - F(1-)$, the atom size at 1. F can thus be represented as $F = a\delta + (1-a)F_c$ where δ is the distribution function with all probability mass at 1 and F_c is the absolutely continuous component of F. The invariant distribution F satisfies

$$\begin{aligned} F(x) &= \int_{[0,1]} P(x|y)\, dF(y) \\ &= aP(x|1) + (1-a)\int_{[0,1)} P(x|y)\, dF_c(y). \end{aligned} \tag{18}$$

Define on $[0,1]$ the improper distribution function $G(x) := (1-a)F_c$, with $G(1) = (1-a)$. Thus, after substitution for $P(\cdot|y)$ in (18), we obtain the integral equation

$$G(x) = ax + \int_0^x \frac{2y(2x^2+x^2y+2x-y)}{x(1+6y+y^2)}\, dG(y) + \int_x^1 \frac{2x(1+y)^2}{1+6y+y^2}\, dG(y).$$

Let the density of G with respect to Lebesgue measure be g. Thus upon differentiation, we obtain for $x \in (0,1]$,

$$g(x) = 1 + a - \frac{2}{x^2}\int_0^x \frac{x^2-y^2}{(1+6y+y^2)} g(y)\, dy. \tag{19}$$

The analytic solution to this equation in terms of elementary functions has eluded us, and, due to the simple and natural formulation of this problem, we see it as an interesting open problem for other researchers. We do have, however, a series solution to (19). Firstly, we see from (19) that

$$g(0+) = 1 + a + \lim_{x\downarrow 0}\frac{xg(x)}{1+6x+x^2} = 1 + a + \lim_{x\downarrow 0} xg(x).$$

Thus, if $g(0+)$ is finite, then $g(0+) = 1 + a$. Suppose for the moment that it is finite and defining $g(0) := g(0+)$, we may attempt a series solution of the form $g(x) = b_0 + b_1x + b_2x^2 + \ldots$ where $b_0 = 1 + a$. This is conveniently done first using the differential equation which follows from the usual manipulations on (19), namely the equation,

$$x(1+6x+x^2)g''(x) + 3(1+6x+x^2)g'(x) + 4g(x) = 0. \tag{20}$$

One finds that $b_1 = -4b_0/3$ and, for $n \geq 2$, that $b_n = -b_{n-2} - 2b_{n-1}(3n^2+6n+2)/(n(n+2)$. Thus

$$g(x) = (1+a)\big[1 - \tfrac{4}{3}x + \tfrac{11}{3}x^2 - \tfrac{112}{9}x^3 + \tfrac{1283}{27}x^4 - \ldots\big]. \tag{21}$$

The coefficients b_n grow in magnitude, for large n, like $(3+2\sqrt{2})^n$; this can be established from the "large n" behaviour of b_n/b_{n-1} in the recurrence for b_n given above. So the radius of convergence for this series is $3 - 2\sqrt{2}$. Substitution back into (19) for $0 < x < 3 - 2\sqrt{2}$ is readily done using the identity, valid within this range,

$$\int_0^x \frac{y^n\, dy}{1+6y+y^2} = \frac{x^{n+1}}{n+1} - 6\frac{x^{n+2}}{n+2} + 35\frac{x^{n+3}}{n+3} - 204\frac{x^{n+4}}{n+4} - 1189\frac{x^{n+5}}{n+5} + \ldots.$$

The coefficients $1, -6, 35, -204, 1189, \ldots$ are generated by $1/(1+6x+x^2)$. Thus by this substitution, we indeed establish that (21) is a solution of (19) for $0 < x <$

$3-2\sqrt{2}$. This series solution is, however, inadequate for our purposes. In particular we cannot evaluate a, which we know is equal to $1-\int_0^1 g(x)dx$, from a form which is undefined on part of $[0,1]$.

We thus turn to series of the form $g(x)=c_0+c_1(1-x)+c_2(1-x)^2+\dots$ and find, from (20), that $c_2=(6c_1-c_0)/4$ and, for $n\geq 2$, that

$$c_n=\frac{(2n-1)c_{n-1}}{n-1}-\frac{(3n-2)(3n-5)c_{n-2}}{8n(n-1)}+\frac{(n-3)c_{n-3}}{8n}.$$

Thus the series is

$$\begin{aligned} g(x)=&c_0+c_1(1-x)+\left(\tfrac{3}{2}c_1-\tfrac{1}{4}c_0\right)(1-x)^2+\left(\tfrac{23}{12}c_1-\tfrac{5}{12}c_0\right)(1-x)^3\\ &+\left(\tfrac{55}{24}c_1-\tfrac{35}{64}c_0\right)(1-x)^4+\left(\tfrac{2537}{960}c_1-\tfrac{79}{120}c_0\right)(1-x)^5+\dots. \end{aligned}$$

We have used the recurrence in c_n and *Mathematica* to compute the exact rational coefficients of c_1 and c_0 in a vast number of successive terms. The ratio of these two coefficients converges quite rapidly to a negative constant. As a result, we have concluded that the coefficients of $(1-x)^n$ converge to zero iff c_1/c_0 equals "minus the reciprocal of that constant", or 0.26426474646613275154584. Such convergence to zero is necessary if a finite value of $g(0)$ is to emerge from the series. Embracing this requirement, we are left with only one unknown entity in the series, namely c_0. There are, however, two additional equations.

$$g(0)=1+a \qquad \int_0^1 g(x)dx=1-a.$$

Adding these together eliminates a and gives an equation which yields c_0 numerically as 0.6795922821356534099481. The constants c_0 and c_1 can be computed to arbitrary accuracy quite easily, but to lower accuracy, we may write the series as

$$\begin{aligned} g(x)=0.6796&+0.1796(1-x)+0.0995(1-x)^2+0.0611(1-x)^3\\ &+0.0399(1-x)^4+0.0272(1-x)^5+\dots. \end{aligned} \tag{22}$$

A slight surprise from our very precise calculations is that $c_1=c_0-\frac{1}{2}$ exactly. Finally, a, the probability of the shape being a square, can be accurately calculated to be 0.16158719822985795896.

Our series (22) is convergent for all $x\in[0,1]$. Thus it is finite at $x=0$, consistent with an assumption made in the argument that it be so. We have not investigated solutions where $g(0)$ is infinite, because we find, through direct substitution into our original integral equation (19), that (22) is a solution. From the obvious ϕ-recurrence of the Markov process, it is the unique invariant solution to our problem and we need look no further. Figure 7(c) plots the solution g.

Whilst our series is a worthwhile solution to our problem, the failure to solve this problem analytically has frustrated us. We have, however, solved a related problem with a different hitting measure. If the congruent shape is rotated with probability $p>0$, fixed for all iterations, with position being uniformly distributed over all possible rotated positions, or is *not* rotated with probability $q := 1 - p$, again with uniform distribution over these possible positions, we can establish that

$$a = \frac{p - 4p^{3/2}I}{1 + p - 4p^{3/2}I},$$

where

$$I = \int_0^1 \frac{y^{\sqrt{p}}\,dy}{(1+y)^2},$$

whilst $g(x) = (1-a)\sqrt{p}x^{\sqrt{p}-1}$. When $p = 1$, the probability mass a at $x = 1$ is $(3 - 4\log 2)/(4 - 4\log 2)$ and, given that the rectangle is not a square, its shape is uniformly distributed on $[0, 1]$. When $p = 0$, shapes degenerate to 0.

References

[1] Blaschke, W. (1936). *Vorlesungen über Integralgeometrie.* Deutsch Verlag Wissenschaft, Berlin.

[2] Chen, F. K. C. and Cowan, R. (1998). Invariant distributions for shapes in sequences of randomly-divided rectangles. *Adv. Appl. Prob.* **30**, ?–?.

[3] Cowan, R. (1997). Shapes of rectangular prisms after repeated random division. *Adv. Appl. Prob.* **29**, 26–37.

[4] Crofton, M. W. (1885). *Probability.* Encyclopedia Britannica, 9th ed., **19**, 768–788, London.

[5] Mannion, D. (1988). A Markov chain of triangle shapes. *Adv. Appl. Prob.* **20**, 348–370.

[6] Mannion, D. (1990). Convergence to collinearity of a sequence of random triangle shapes. *Adv. Appl. Prob.* **22**, 831–844.

[7] Kendall, D. G. (1977). The diffusion of shape. *Adv. Appl. Prob.* **9**, 428–430.

[8] Rényi, A. and Sulanke, R. (1963). Über die konvexe Hüll von n zufällig gewählten Punkten. *Z. Wahr.* **2**, 75–84.